

Establishing the UK DNA Bank for motor neuron disease (MND)

Smith, Lucy; Cupid, B. C.; Dickie, B. G M; Al-Chalabi, A.; Morrison, K. E.; Shaw, C. E.; Shaw, P. J.

DOI:

[10.1186/s12863-015-0236-6](https://doi.org/10.1186/s12863-015-0236-6)

License:

Creative Commons: Attribution (CC BY)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Smith, L, Cupid, BC, Dickie, BGM, Al-Chalabi, A, Morrison, KE, Shaw, CE & Shaw, PJ 2015, 'Establishing the UK DNA Bank for motor neuron disease (MND)', *BMC Genetics*, vol. 16, no. 1, 84.

<https://doi.org/10.1186/s12863-015-0236-6>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Eligibility for repository: Checked on 14/12/2015

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Manuscript Number:		
Full Title:	Establishing the UK DNA bank for motor neuron disease (MND)	
Article Type:	Methodology article	
Section/Category:	Complex traits and quantitative genetics	
Funding Information:	Wellcome Trust (070122/A/02/Z)	Dame Professor Pamela J Shaw
	MND Association (Shaw/Nov02/6700-3)	Dame Professor Pamela J Shaw
Abstract:	<p>In 2003 the Motor Neurone Disease (MND) Association, together with The Wellcome Trust, funded the creation of a national DNA Bank specific for MND. It was anticipated that the DNA Bank would constitute an important resource to researchers worldwide and significantly increase activity in MND genetic research. The DNA Bank houses over 3000 high quality DNA samples, all of which were donated by people living with MND, family members and non-related controls, accompanied by clinical phenotype data about the patients. Today the primary focus of the UK MND DNA Bank still remains to identify causative and disease modifying factors for this devastating disease.</p>	
Corresponding Author:	Lucy Smith, Ph.D MND Association Northampton, UNITED KINGDOM	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	MND Association	
Corresponding Author's Secondary Institution:		
First Author:	Lucy Smith, Ph.D	
First Author Secondary Information:		
Order of Authors:	Lucy Smith, Ph.D Belinda Cupid Brian GM Dickie Ammar Al-Chalabi Christopher E Shaw Karen E Morrison Pamela J Shaw	
Order of Authors Secondary Information:		
Manuscript Classifications:	10.030: Epigenetics and chromosome biology; 10.040: Functional genetics; 20.100: Functional genomics; 20.120: Genetic screens; 20.340: Whole genome sequencing; 40.020: Genetic basis of human disease	
Opposed Reviewers:	Garth Nicholson, Professor of medicine ANZAC Research Institute garth.nicholson@sydney.edu.au expert in the field, no competing interests with any author	
	Markus Weber, Professor of Neurology Head Neuromuscular Diseases Unit, ALS Clinic Kantonsspital St.Gallen markus.weber@kssg.ch	

	expert in the field No competing interest
	Robert Swingler, Medicine Consultant, Ninewells Hospital, NHS Tayside, Dundee
	Expert in the field no conflict of interest
	Jochen Weishaupt, Professor Universitat Ulm jochen.weishaupt@uni-ulm.de expert in the field no conflict of interest

Establishing the UK DNA Bank for motor neuron disease (MND)

*Smith L¹, Cupid BC¹, Dickie BGM¹, #Al-Chalabi A², #Morrison KE³, #Shaw CE², #Shaw PJ⁴

*corresponding author lucy.smith@mndassociation.org

#Principal Investigators for the DNA Bank, listed in alphabetical order.

¹Motor Neurone Disease Association, PO Box 246, Northampton, NN1 2PR, UK;

belinda.cupid@mndassociation.org; brian.dickie@mndassociation.org

²NIHR Biomedical Research Unit in Dementia, Department of Clinical Neuroscience, King's College London, London, SE5 8AF, UK; ammar.al-chalabi@kcl.ac.uk; chris.shaw@kcl.ac.uk

³Institute of Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, UK; k.morrison@bham.ac.uk

⁴Sheffield Institute for Translational Neuroscience, University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ, UK; pamela.shaw@sheffield.ac.uk

The authors declare that they have no competing interests.

Abstract

In 2003 the Motor Neurone Disease (MND) Association, together with The Wellcome Trust, funded the creation of a national DNA Bank specific for MND. It was anticipated that the DNA Bank would constitute an important resource to researchers worldwide and significantly increase activity in MND genetic research. The DNA Bank houses over 3000 high quality DNA samples, all of which were donated by people living with MND, family members and non-related controls, accompanied by clinical phenotype data about the patients. Today the primary focus of the UK MND DNA Bank still remains to identify causative and disease modifying factors for this devastating disease.

Keywords: Motor Neurone Disease (MND), Amyotrophic Lateral Sclerosis (ALS), Biobank.

Motor Neuron Disease (MND) is a fatal, rapidly progressive disease that affects the brain and spinal cord and which ultimately leads to respiratory failure around 2-5 years following symptom onset [1, 2]. Approximately 1 in 300 people develop MND but its prevalence is low, at about 6-8 in 100,000 because of short life expectancy [3]. There is no diagnostic test and treatment is largely palliative, with only one agent, riluzole, having a modest effect in extending survival. Genetic factors undoubtedly play a role in most cases of the disease, both in pathogenesis and rate of progression, with about 5-10% of all patients having a clear family history of MND and in some cases, frontotemporal dementia [4, 5, 6]. Over 100 genes have now been implicated in the causation of MND [7]. No consistent environmental risk factor has been identified, although it is possible that such factors may trigger disease in genetically susceptible individuals, and therefore it is plausible that apparent sporadic cases of MND will be genetically determined to some degree [8, 9].

An essential starting point for successful genetic research is access to high quality samples, accompanied by detailed clinical information. Large-scale gene sequencing and association studies need many thousands of samples to be screened such that results are statistically significant. Access to such samples had become a major obstacle in exploring the pathogenesis of MND and the concept of a MND DNA Bank was born. The objectives of the initial study were threefold: 1) To collect cohorts of patient, parent/sibling and control samples from sporadic and familial MND; 2) To collect clinical information in order to examine susceptibility traits in clinical subgroups of MND; 3) To make this resource available to the international research community and to foster collaboration between research teams, in order to identify genetic risk factors for MND.

ORGANISATIONAL STRUCTURE OF THE UK MND DNA BANK

1 The UK MND DNA Bank was a collaborative project adopting a 'Hub and Spoke' model, with three
2 regional 'Hub' centres linking with a total of 16 'Spoke' centres (Table 1). Samples were obtained
3
4 from sporadic and familial MND patients attending MND clinics in the UK, their spouses (or other
5
6 genetically unrelated controls) and blood relatives. Samples within the UK MND DNA Bank are
7
8 housed at CIGMR Biobank, at the University of Manchester. In addition, as one of the Public Health
9
10 England collections, the European Collection of Cell Cultures (ECACC) manages the transformation
11
12 and storage of EBV-transformed lymphocytes derived from blood samples from participants
13
14 providing an everlasting supply of DNA for the Bank.
15
16
17
18
19
20

21 The MND Association's Biomedical Research Advisory Panel (BRAP) oversee the governance and
22
23 the strategic development of the DNA Bank, ensuring that samples are utilised in an appropriate
24
25 fashion, and that any clinical information requested is appropriate for the proposed study. The
26
27 Technical Access Committee (TAC) at CIGMR Biobank, determine sample requirements for the
28
29 technology platform to be used, the quantity of sample required and ensure any leftover samples are
30
31 returned or destroyed. All applications for access to the samples are judged on merit. In order to
32
33 receive material and clinical information from the DNA Bank, all applicants must agree to the terms
34
35 and conditions of sample use (see supplementary info). This specifies the user and specific purpose
36
37 for which the samples and data are to be licensed, including standard terms as to the ownership,
38
39 exploitation and dissemination of results, and requirements that the user conforms to the terms of the
40
41 participants' consent.
42
43
44
45
46
47
48
49
50

51 **SAMPLE COLLECTION, STORAGE AND QUALITY CONTROL**

52

53 Sample collection began in 2003. All participants were over 18 years of age. In order to ensure that
54
55 the patient cohort was representative of disease prognosis, patients must have experienced symptom
56
57 onset (significant muscle weakness) on or after January 2002. All patients fulfilled El Escorial
58
59
60
61
62
63
64
65

1 criteria for probable or definite Amyotrophic Lateral Sclerosis (ALS) [10]. Patients presenting with
2 Progressive Muscular Atrophy (PMA), Primary Lateral Sclerosis (PLS) or Progressive Bulbar Palsy
3 (PBP) were also included in the study. Patients were recruited by consultant neurologists with a
4 specialist interest in MND in participating centres. Patients participating in other clinical research
5 projects were not excluded from the study. Blood samples were also collected from consenting
6 partners/carers, providing some degree of matching in terms of age, education, environmental
7 exposure and often ethnicity. Where patients presented with familial MND, blood samples were
8 collected from family members for linkage analysis. Where patients presented with sporadic MND,
9 where possible, blood samples were also collected from parents or from a parent and sibling, to give
10 so-called 'Trio samples' increasing the amount of genetic information available for researchers.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Informed consent to participate was sought from all patients, family members and controls. Ethical
27 approval for the collection of samples and the creation of the UK MND DNA Bank was given by the
28 Trent Research Ethics Committee in February 2003 ref MREC/02/4/107 and in July 2009, ref
29 09/HO405/32. Participants were provided with detailed information and contact details and could
30 withdraw from the study at any time. The samples were pseudo-anonymised and an online clinical
31 database was developed to facilitate data entry and collection by the research nurses and enable
32 tracking of trends in clinical parameters such as symptom onset and presentation for data analysis.
33 Storage and access to this data set is in accordance with the UK Data Protection Act 1998 [11].
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Prior to 2010, DNA extraction from donated blood samples was carried out at individual Hub centres
50 using the Nucleon BACC3 protocol (Amersham, UK). Extracted DNA was sent to CIGMR Biobank
51 for long-term storage. On receipt, all DNA samples were run on 1% agarose gels alongside
52 molecular weight markers of appropriate size to check integrity. From August 2010, DNA extraction
53 was carried out at CIGMR Biobank using automated robotic processing under ISO900:2000
54
55
56
57
58
59
60
61
62
63
64
65

operating standards. In all cases, both when imported from Hub centres, or extracted by CIGMR Biobank themselves, DNA concentration was measured using a nanodrop spectrophotometer. Samples with OD ratios outside the normal range were removed from the cohort and contaminants washed using ethanol precipitation. Final DNA concentration was measured using Quant-iT™ Picogreen® dsDNA Assay Kit (Invitrogen™ Life Technologies, UK). Samples were measured on 96 well plates, in triplicate against standards of known concentration for quality control. DNA aliquots are stored in 2D bar coded tubes for sample tracking purposes. A relational database recorded the 2D barcodes associated with each patient/donor ID. All samples within the collection were screened for gender using PCR on presumed duplicate samples according to standard protocols. Samples with a mismatch between the expected gender as recorded in the patient information, and actual gender as confirmed by PCR, were rescreened using an alternative PCR method of gender identification based on the absence/presence of *Alu* sequence [12]. Any samples with a confirmed discrepancy were ring fenced from the collection and suspended from the in-house laboratory management system.

Peripheral blood lymphocytes (PBLs) were isolated from whole blood samples at ECACC using density gradient centrifugation. An aliquot of untransformed PBLs was stored in liquid nitrogen for safekeeping, whilst the remaining PBLs were transformed using the Epstein Barr virus according to standard protocols [13]. The resulting lymphoblastoid cell lines were cryopreserved and are used to restock the DNA Bank when stock levels become low.

4. THE UK MND DNA BANK

In October 2012, at the end of the collection period, the UK MND DNA Bank comprised 3159 high quality DNA samples. Of these 1344 samples were taken from individuals diagnosed with sporadic MND (see figure 1A and 1B). There were 133 familial MND samples within the collection and a

1 further 500 samples taken from family members, including samples that form 28 parent trio sets and
2 27 sibling trio sets. The remaining 1085 samples were taken from controls. In line with population-
3 based demographic for the disease [14] the breakdown of gender in the collection is around 60%
4 male (Figure 1A). The average age of onset was approximately 62 years of age (Figure 1C). Each
5 sample is accompanied by a minimum dataset of: age at which the samples were taken; gender;
6 disease status; and where appropriate diagnostic certainty (El Escorial Status) and age of onset
7 (calculated from date of birth and date of symptom onset). An extended dataset has been collected
8 for as many participants as possible but it is not a complete dataset for the entire collection (see Table
9 2). In total 2653 frozen lymphoblastoid cell lines are held in storage at ECACC following a PBL
10 transformation success rate of 97%. Of these, 1267 samples were generated from whole blood taken
11 from patients with sporadic MND. 115 cell lines were generated from familial samples and the
12 remaining 1058 cells lines have been established using blood samples obtained from control or
13 family members (see Figure 1D).

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34 Table 3 shows the success rates for PCRs performed on the DNA samples within the collection. The
35 failure rate of the quality control assay was less than 1.5% suggesting that the quality of DNA within
36 the collection is very high. The gender results from these assays were directly compared to the
37 gender recorded for individuals on the clinical database. Where there was a discrepancy between the
38 expected gender and that determined in the assay, patient clinical notes were rechecked. In the
39 absence of a clerical error, samples were rescreened using both the original AMEL marker and an
40 alternative gender marker, the Human ALU expansion [12]. Sixty eight samples that continued to
41 show a discrepancy between the expected gender and the assay gender were ring fenced from the
42 collection and suspended from the laboratory management system.

The UK MND DNA Bank was designed to be available to the international research community. DNA samples from the bank represent an incident not prevalent population and are unlikely to be biased. The fundamental guarantee that any DNA Bank must be able to give is that it can provide high quality DNA samples with good integrity and accompanying high quality clinical data. Understandably, a constraint of the DNA Bank is that the genomic DNA supply itself is limited and although cell lines have been established, the DNA from such cell lines may have sequence changes compared with the original genomic samples. This fact must be considered when choosing to use cell line derived DNA even if the DNA itself is of a high standard as demonstrated by the rigorous quality control assays in place. As part of the governance of the DNA Bank, the MND Association must ensure compliance with legal and regulatory requirements. The Association must also guarantee that the resource adheres to rigorous research standards and is used in the further understanding of motor neuron disease, this includes prioritising access to those parts of the DNA Bank that are limited in availability, clarifying intellectual property rights and disseminating the results that flow from it.

To date more than twenty projects have withdrawn samples from the DNA Bank. DNA samples have been used in complex, technical protocols such as genotyping, gene sequencing and genome-wide association studies and numerous papers have been published or are in press [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 and 28]. Importantly, projects using samples from the DNA Bank have directly led to the detection of several MND causing genes including *C9orf72* and more recently *Tub4A* [15, 17, 22, 23 and 28]. With researchers now encouraged to publish in an open access format as part of the DNA Bank governance, and to deposit data from sequencing projects within accessible databases such as ALSOD: the Amyotrophic Lateral Sclerosis Online Database [7] and European Genome-Phenome archive [29], the dissemination and discussion of results by the research community is ensured. In 2014 a proposal to perform whole genome sequencing on DNA samples from the UK MND DNA Bank as part of the international collaboration called Project MinE [30]

1 was approved. This exciting project will allow Next Generation Sequencing data to be collected from
2 DNA Bank samples and shared across research groups. The data will also confirm the accuracy of
3
4 existing studies through imputation. It is hoped that sequencing DNA Bank samples will allow the
5
6 identification of rare variants responsible for sporadic disease, continuously widening our knowledge
7
8 about how genetic changes can contribute to MND.
9
10

11
12
13
14
15 Ethical approval to extend the use of the cell lines beyond their original scope of providing an
16
17 everlasting supply of DNA was granted in 2014 by the Derby- East Midlands Research Ethics
18
19 Committee ref no. 14/EM/1088. This change in permission will potentially allow researchers to
20
21 generate primary neuronal cultures and highly desirable induced pluripotent stem (iPS) cell lines
22
23 from the cell lines stored at ECACC. The iPS cell lines could act as new disease models for drug
24
25 screening and other potential treatments, as well as acting as tools for analysing downstream
26
27 mechanisms involved in disease pathogenesis. Clearly the role of the DNA Bank in the governance
28
29 of such samples will be paramount; it is simply not enough to provide high quality samples, but
30
31 following how those samples have been used and ensuring the results are disseminated and discussed
32
33 is the only way to ensure research continues to move forward. The original scope of the DNA Bank
34
35 was to make a quantal difference in our understanding of MND and it is well on the way to fulfilling
36
37 this promise.
38
39
40
41
42
43
44
45

46 **Acknowledgments:** This project was supported by a programme grant from the Motor Neurone Disease
47
48 Association (grant ref 6700). Information regarding the DNA Bank, including the terms and conditions for
49
50 sample use can be accessed on the MND Association website [31]. The creation of the lymphoblastoid cell lines
51
52 was supported by a grant from the Wellcome Trust (grant ref 070122/A/02/Z). Ethical approval for the
53
54 creation of the UK MND DNA Bank was given by the Trent Research Ethics Committee in February 2003 ref
55
56 MREC/02/4/107 and in July 2009, ref 09/HO405/32. Subsequent approval for the extended use of the cell
57
58 lines within the DNA Bank was approved by the Derby- East Midlands Research Ethics Committee in Sept
59
60
61
62
63
64
65

2014, ref no. 14/EM/1088. We are grateful for the support of the Dementias and Neurodegenerative Disease Research Network (DeNDRoN). We would like to acknowledge our partners at the CIGMR Biobank (formerly BioBanking Solutions (BBS), UK DNA Bank Network (UDBN)) based at the University of Manchester and the European Cell Culture Collection at Public Health England. This project would now have been possible without the support of collaborators in the participating centres across the UK [31]. Thank you to all the people with MND and their families who participated in this project.

References:

1. McDermott CJ and Shaw PJ. Diagnosis and management of motor neurone disease. *BMJ* 2008; 336: 658-662. doi: 10.1136/bmj.39493.511759.BE.
2. Chio A, Mora G, Calvo A, Mazzini L, Bottacchi E, Mutani R. Epidemiology of ALS in Italy: a 10-year prospective study. *Neurology* 2009; 72: 725-731. doi: 10.1212/01.wnl.0000343008.26874.d1.
3. Johnston CA, Stanton BR, Turner MR, Gray R, Blunt AH, Butt D, Ampong MA, Shaw CE, Leigh PN, Al-Chalabi A. ALS in an urban setting: a population based study of inner city London J. *Neurology* 2006; 253: 1642-43.
4. Mackenzie IR, Rademakers R, Neumann M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurology* 2010; 9: 995-1007.
5. Byrne S, Walsh C, Lynch C, Bede P, Elamin M, Kenna K, McLaughlin R, Hardiman O. Rate of familial ALS: a systematic review and metanalysis. *J. Neurol Neurosurg Psychiatry* 2011; 82: 623-627. doi: 10.1136/jnnp.2010.224501.
6. Byrne S, Heverin M, Elamin M, Bede P, Lynch C, Kenna K, Maclaughlin R, Walsh C, Al-Chalabi A, Hardiman O. Aggregation of neurologic and neuropsychiatric disease in amyotrophic lateral sclerosis kindreds: A population based case-control cohort study of familial and sporadic amyotrophic lateral sclerosis. *Ann Neurol* 2013; 74: 699-708. doi: 10.1002/ana.23969.
7. Lill CM, Abel O, Bertram L, Al-Chalabi A. Keeping up with the genetic discoveries in ALS: The ALSoD and ALSGene database. *ALS* 2011; 12: 238-249. doi: 10.3109/17482968.2011.584629.
8. Al-Chalabi A, Lewis CM. Modelling the effects of penetrance and family size on rates of sporadic and familial disease. *Hum Hered* 2011; 71: 281-88. doi: 10.1159/000330167.

9. Factor-Litvak P, Al-Chalabi A, Ascherio A, Bradley W, Chío A, Garruto R, Hardiman O, Kamel F, Kasarskis E, McKee A, Nakano I, Nelson LM, Eisen A. Current pathways for epidemiological research in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener.* 2013;14 Suppl 1:33-43. doi: 10.3109/21678421.2013.778565.
10. Belsh JM. ALS diagnostic criteria of El Escorial Revisited: do they meet the needs of clinicians as well as researchers? *Amyotroph Lateral Scler Other Motor Neuron Disord.* 2000 Mar;1 Suppl 1:S57-60.
11. The Data Protection Act 1998, National Archives, <http://www.legislation.gov.uk/ukpga/1998/29/contents>.
12. Hedges DJ, Walker JA, Callinan PA, Shewale JG, Sinha SK, Batzer MA. Mobile element-based assay for human gender determination. *Analytical Biochemistry* 2003; 312: 77–79
13. Blick D, Cooper J, Baker N, Bracegirdle P, Biggins J, Burnet W. Generation of cell lines using Epstein-Barr Virus (EBV) transformation of small volumes of cryo-preserved whole blood and the use of bench-top flow cytometry to achieve high and reproducible success rates. http://www.nature.com/app_notes/nmeth/2011/111312/pdf/an8194.pdf
14. Alonso A, Logroscino G, Jick SS and Hernan MA. Incidence and lifetime risk of motor neuron disease in the United Kingdom: a population based study. *Eur J Neurol* 2009; 16: 745-51.
15. Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, Johnson L, Veldink JH, van Es MA, van den Berg LH, Robberecht W, Van Damme P, Hardiman O, Farmer AE, Lewis CM, Butler AW, Abel O, Andersen PM, Fogh I, Silani V, Chiò A, Traynor BJ, Melki J, Meininger V, Landers JE, McGuffin P, Glass JD, Pall H, Leigh PN, Hardy J, et al. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol.* 2010 Oct; 9(10):986-94. doi: 10.1016/S1474-4422(10)70197-6
16. Smith BN, Newhouse S, Shatunov A, Vance C, Topp S, Johnson L, Miller J, Lee Y, Troakes C, Scott KM, Jones A, Gray I, Wright J, Hortobágyi T, Al-Sarraj S, Rogelj B, Powell J, Lupton M, Lovestone S, Sapp PC, Weber M, Nestor PJ, Schelhaas HJ, Asbroek AA, Silani V, Gellera C, Taroni F, Ticozzi N, Van den Berg L, Veldink J, et al. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet.* 2013 Jan; 21(1):102-8. doi: 10.1038/ejhg.2012.98.

17. Van Rheenen W, Diekstra FP, van Doormaal PT, Seelen M, Kenna K, McLaughlin R, Shatunov A, Czell D, van Es MA, van Vught PW, van Damme P, Smith BN, Waibel S, Schelhaas HJ, van der Kooi AJ, de Visser M, Weber M, Robberecht W, Hardiman O, Shaw PJ, Shaw CE, Morrison KE, Al-Chalabi A, Andersen PM, Ludolph AC, Veldink JH, van den Berg LH. H63D polymorphism in HFE is not associated with amyotrophic lateral sclerosis. *Neurobiol Aging*. 2013 May; 34(5):1517.e5-7. doi: 10.1016/j.neurobiolaging.2012.07.020.
18. Jones AR, Woollacott I, Shatunov A, Cooper-Knock J, Buchman V, Sproviero W, Smith B, Scott KM, Balendra R, Abel O, McGuffin P, Ellis CM, Shaw PJ, Morrison KE, Farmer A, Lewis CM, Leigh PN, Shaw CE, Powell JF, Al-Chalabi A. Residual association at C9orf72 suggests and alternative ALS causing hexanucleotide repeat. *Neurobiol of Aging* 2013; 34(9): 2234.e1-7. doi: 10.1016/j.neurobiolaging.2013.03.003.
19. Mok K, Laaksovirta H, Tienari PJ, Peuralinna T, Myllykangas L, Chiò A, Traynor BJ, Nalls MA, Gurunlian N, Shatunov A, Restagno G, Mora G, Nigel Leigh P, Shaw CE, Morrison KE, Shaw PJ, Al-Chalabi A, Hardy J, Orrell RW. Homozygosity analysis in amyotrophic lateral sclerosis. *Eur J Hum Genet*. 2013; 12:1429-35. doi: 10.1038/ejhg.2013.59.
20. Buchman VL, Cooper-Knock J, Connor-Robson N, Higginbottom A, Kirby J, Razinskaya OD, Ninkina N, Shaw PJ. (2013) Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridization. *Mol Neurodegener*. 2013 Apr 8; 8: 12. doi: 10.1186/1750-1326-8-12.
21. Ismail A, Cooper-Knock J, Highley JR, Milano A, Kirby J, Goodall E, Lowe J, Scott I, Constantinescu CS, Walters SJ, Price S, McDermott CJ, Sawcer S, Compston DA, Sharrack B, Shaw PJ. Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of C9ORF72. *J Neurol Neurosurg Psychiatry*. 2013; 84(1): 79-87. doi: 10.1136/jnnp-2012-303326.
22. Cooper-Knock J, Higginbottom A, Connor-Robson N, Bayatti N, Bury JJ, Kirby J, Ninkina N, Buchman VL, Shaw PJ. C9ORF72 transcription in frontotemporal dementia case with two expanded alleles. *Neurology*. 2013; 81(19):1719-21. doi: 10.1212/01.wnl.0000435295.41974.2e.
23. Fogh I, Ratti A, Gellera C, Lin K, Tiloca C, Moskvina V, Corrado L, Sorarù G, Cereda C, Corti S, Gentilini D, Calini D, Castellotti B, Mazzini L, Querin G, Gagliardi S, Del Bo R, Conforti FL, Siciliano G, Inghilleri M, Saccà F, Bongioanni P, Penco S, Corbo M, Sorbi S, Filosto M, Ferlini A, Di Blasio AM,

- Signorini S, Shatunov A, et al. A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. *Hum Mol Genet* 23(8): 2220-31. doi: 10.1093/hmg/ddt587.
24. Goris A, van Setten J, Diekstra F, Ripke S, Patsopoulos NA, Sawcer SJ; International Multiple Sclerosis Genetics Consortium, van Es M; Australia and New Zealand MS Genetics Consortium, Andersen PM, Melki J, Meininger V, Hardiman O, Landers JE, Brown RH Jr, Shatunov A, Leigh N, Al-Chalabi A, Shaw CE, Traynor BJ, Chiò A, Restagno G, Mora G, Ophoff RA, Oksenberg JR, Van Damme P, Compston A, Robberecht W, Dubois B, van den Berg LH, et al. No evidence for shared genetic basis of common variants in multiple sclerosis and amyotrophic lateral sclerosis. *Hum Mol Genet.* 2014; 23(7):1916-22.doi: 10.1093/hmg/ddt574.
25. Diekstra FP, Van Deerlin VM, van Swieten JC, Al-Chalabi A, Ludolph AC, Weishaupt JH, Hardiman O, Landers JE, Brown RH Jr, van Es MA, Pasterkamp RJ, Koppers M, Andersen PM, Estrada K, Rivadeneira F, Hofman A, Uitterlinden AG, van Damme P, Melki J, Meininger V, Shatunov A, Shaw CE, Leigh PN, Shaw PJ, Morrison KE, Fogh I, Chiò A, Traynor BJ, Czell D, Weber M, et al. C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: A genome-wide meta-analysis *Ann Neurol* 76(1):120-33. doi: 10.1002/ana.24198.
26. Savage AL, Wilm TP, Khursheed K, Shatunov A, Morrison KE Shaw PJ, Shaw CE, Smith B, Breen G, Al-Chalabi A, Moss D, Bubb VJ, Quinn JP. An evaluation of a SVA retrotransposon in the FUS promoter as a transcriptional regulator and its association to ALS. *PLoS One.* 2014; 9(6): e90833. doi: 10.1371/journal.pone.0090833. eCollection 2014.
27. Cooper-Knock J, Walsh MJ, Higginbottom A, Robin Highley J, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* 2014 Jul; 137 (Pt 7):2040-51. doi: 10.1093/brain/awu120.
28. Smith BN, Ticozzi N, Fallini C, Gkazi AS, Topp S, Kenna KP, Scotter EL, Kost J, Keagle P, Miller JW, Calini D, Vance C, Danielson EW, Troakes C, Tiloca C, Al-Sarraj S, Lewis EA, King A, Colombrita C, Pensato V, Castellotti B, de Bellerocche J, Baas F, ten Asbroek AL, Sapp PC, McKenna-Yasek D, McLaughlin

RL, Polak M, Asress S, Esteban-Pérez J, et al. Exome-wide Rare Variant Analysis Identifies TUBA4A Mutations Associated with Familial ALS. *Neuron*. 2014 Oct 22; 84(2) :324-31. doi: 10.1016/j.neuron.2014.09.027. Epub 2014 Oct 22.

29. The European Genome-Phenome archive website: <https://www.ebi.ac.uk/ega/home>

30. The Project MinE website: <https://www.projectmine.com/>

31. The MND Association Website: <http://www.mndassociation.org/dnabank>

FIGURE LEGENDS

Table 1: Hub and Spoke Model for Sample Collection

The UK MND DNA Bank was a collaborative project adopting a ‘hub and spoke’ model with three regional ‘hub centres’ linking with a total of 16 ‘spoke centres’. These included hospitals that are part of the MND Association’s Care Centre Network and centres which form part of the Department of Health / NIHR Dementias and Neurodegenerative Diseases Research Network (DeNDRoN). The three Hub Centres were established at London (King’s College Hospital), Sheffield (Royal Hallamshire Hospital), and Birmingham (Queen Elizabeth Hospital). Recruitment to the study and sample collection was coordinated at the Hub centres by a DNA Bank Co-ordinator based in London. Samples were obtained from sporadic and familial MND patients attending MND clinics in the UK, their spouses (or other genetically unrelated controls) and blood relatives. A DNA Bank research nurse was affiliated to each Hub centre to act as patient liaison, collect clinical information from patients, controls and family members, and take blood samples. Laboratory Technicians were also employed at each Hub Centre to assist in the preparation of sample collection packs for participating satellite centres, DNA extraction and final sample storage.

Figure 1: The UK MND DNA Bank

1 The UK MND DNA Bank comprises 3159 high quality DNA samples. Of these 1344 samples were
2 taken from individuals diagnosed with sporadic MND (see figure 1A and 1B). There were 133
3 familial MND samples within the collection and a further 500 samples taken from family members,
4 including samples that form 28 parent trio sets and 27 sibling trio sets. The remaining 1085 samples
5 were taken from controls. In line with previous findings, where MND has been diagnosed, the
6 breakdown of gender in the collection is around 60% male (Figure 1A). The average age of onset
7 was approximately 62 years of age (Figure 1C). In total 2653 frozen lymphoblastoid cell lines are
8 held in storage at ECACC. Of these 1267 samples were generated from whole blood taken from
9 patients with sporadic MND. 115 cell lines were generated from familial samples and the remaining
10 1058 cells lines have been established using blood samples obtained from control or family members
11 (see Figure 1D).

31 **Table 2: Clinical information available from the UK MND DNA Bank**

32 Each sample withdrawn from the UK MND DNA Bank is accompanied by a minimum dataset of:
33 age at which the samples were taken; gender; disease status; and where appropriate diagnostic
34 certainty (El Escorial Status) and age of onset (calculated from date of birth and date of symptom
35 onset). An extended dataset has been collected for as many participants as possible but it is not a
36 complete dataset for the entire collection. The clinical information was collected by the Research
37 Nurse using a brief clinical questionnaire. Identifying data was kept at each Hub centre in secure
38 locations in accordance with the Data Protection Act.

53 **Table 3: Quality Control PCR Assay fail rate**

54 Quality control assays were carried out across the collection. In an initial analysis of the collection
55 768 samples were screened using the ABI Identifiler kit. Following this, a further 2170 samples were

1 screened in a gender-based assay using the AMEL marker. Only 62 samples showed a continued

2 discrepancy between the gender of the actual DNA sample and that stated in the clinical notes.

3
4 Whilst this is still a low level of error for a collection of this size, the samples were ring fenced from
5
6
7 the collection.
8
9

Table 1 : Hub and Spoke Model of Sample Collection

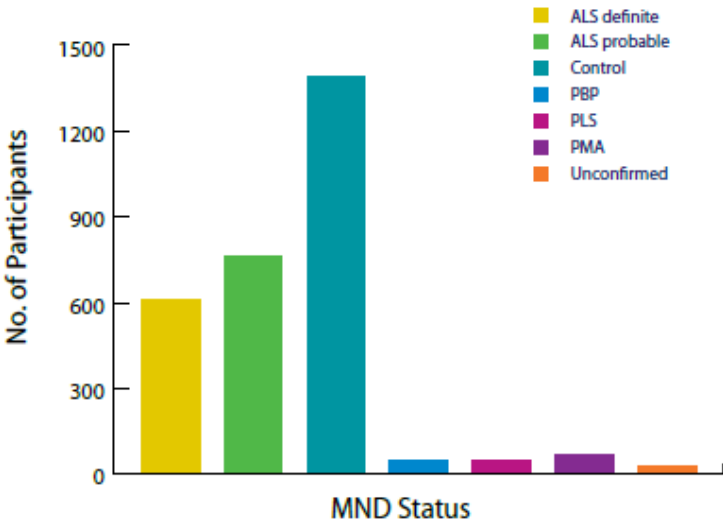
King's College Hospital, London	Queen Elizabeth Hospital, Birmingham	Royal Hallamshire Hospital, Sheffield
Royal Free Hospital London Bart's and The London NHS trust Poole NHS trust, Poole Cambridge University Hospital, Derriford Hospital, Plymouth Southampton University Hospital Queen's Hospital, Romford	John Radcliffe, Oxford Belfast City Hospital, Belfast Walton Neurological Centre, Liverpool Southmead Hospital, Bristol	Royal Preston Hospital Greater Manchester Medical Centre, Manchester Ninewells Hospital, Dundee Queen's Medical Centre, Nottingham Royal Victoria Infirmary, Newcastle

Figure 1: The UK MND DNA Bank Collection

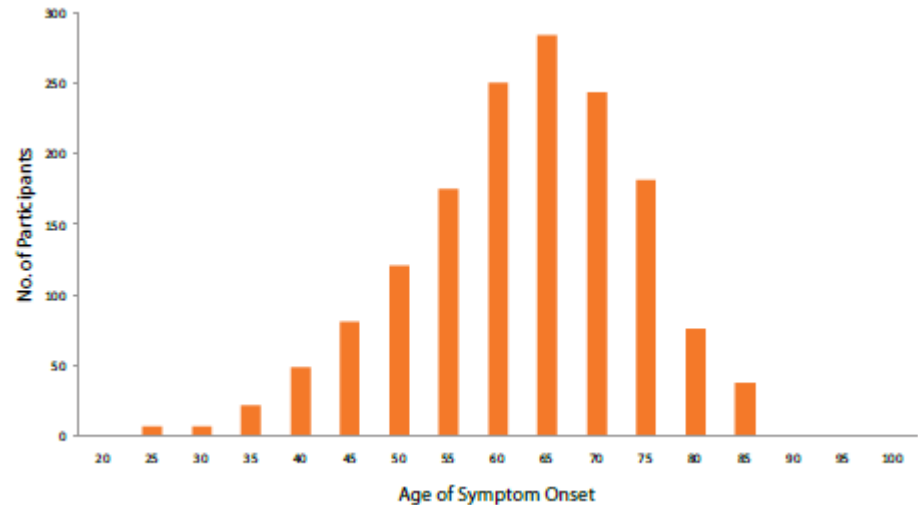
1A: MND status and gender of samples held in the DNA bank

Status	Female	Male	Total numbers
Sporadic	530	834	1,364
Familial	53	80	133
Family*	290	210	500
Unknown	26	46	72
Control	677	408	1,085
Total number	1,576	1,578	3,154

1B: Diagnostic certainty of samples within the DNA bank



1C: Age range of symptom onset within the DNA bank



1D: MND status and gender of cell line samples held at ECACC

Status	Female	Male	Total numbers
Sporadic	489	798	1,287
Familial	44	71	115
Family*	96	79	175
Unknown	25	40	65
Control	627	366	993
Total number	1,281	1,354	2,635

Table 2 : Clinical information available from the UK MND DNA bank

Minimum Dataset

All approved users of the samples will receive a **minimum dataset** of anonymised information on participants.

The fields within the minimum dataset include:

- Age
- Gender
- Affection status (control, familial ALS, sporadic ALS etc)
- Diagnostic certainty (El Escorial status)
- Age of onset

Extended Clinical Dataset

An **extended dataset** has been collected from as many participants as possible. The extended dataset is only available to researchers who collaborate with the Principal Investigators of the DNA Bank.

The fields within the extended dataset include:

- Clinical history (including site of presentation, dominant hand, family history)
- Family tree pedigree
- Investigation and results (including EMG, MRI, Nerve conduction studies)
- Medications (including riluzole)
- ALSFRS score
- Physical examination history (including % FVC, % VC)
- MRC scores

Table 3 : Quality Control PCR fail rates

Type of Assay	No. samples screened	% Assay fail
Abi Identifilier Kit - AMEL Marker	768	1.30
Gender based PCR - AMEL marker	2750	0.62

Total number of individuals screened to confirm gender	% Gender error across collection
3415	1.82

